

## Southwestern Association of Naturalists

---

Vegetation-Index Models Predict Areas Vulnerable to Purple Loosestrife (*Lythrum salicaria*)  
Invasion in Kansas

Author(s): Robert P. Anderson, A. Townsend Peterson and Stephen L. Egbert

Source: *The Southwestern Naturalist*, Vol. 51, No. 4 (Dec., 2006), pp. 471-480

Published by: [Southwestern Association of Naturalists](#)

Stable URL: <http://www.jstor.org/stable/20424756>

Accessed: 07-08-2014 20:18 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at  
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*Southwestern Association of Naturalists* is collaborating with JSTOR to digitize, preserve and extend access to *The Southwestern Naturalist*.

<http://www.jstor.org>

VEGETATION-INDEX MODELS PREDICT AREAS VULNERABLE TO PURPLE LOOSESTRIFE (*LYTHRUM SALICARIA*) INVASION IN KANSAS

ROBERT P. ANDERSON,\* A. TOWNSEND PETERSON, AND STEPHEN L. EGBERT

*Department of Biology, City College of the City University of New York, New York, NY 10031 (RPA)**Natural History Museum, University of Kansas, Lawrence, KS 66045 (ATP)**Department of Geography and Kansas Applied Remote Sensing Program, University of Kansas, Lawrence, KS 66045 (SLE)**\*Correspondent: anderson@sci.cuny.cuny.edu*

**ABSTRACT**—Purple loosestrife (*Lythrum salicaria*) constitutes an invasive species detrimental to wetland habitats in North America. To estimate areas vulnerable to it in Kansas, we modeled the potential geographic distribution of the species by using current records in the state, remotely sensed vegetation-index data from the Moderate Resolution Imaging Spectrometer (MODIS), and the Genetic Algorithm for Rule-Set Prediction (GARP). Models built using only localities from northeastern Kansas (the origin of invasion within the state) consistently predicted test localities in other parts of the state with negligible omission. An additional analysis using records from all counties where the species is known showed a similar prediction. All models indicated suitable conditions for purple loosestrife in much of eastern and central Kansas, as well as in riverine and irrigated areas in the western part of the state. The approach presented here might prove useful for assessing the regional colonization potential of other newly detected invasive species before other studies can be undertaken.

**RESUMEN**—*Lythrum salicaria* constituye una especie invasora perjudicial para los hábitats pantanosos en Norteamérica. Para estimar cuáles son las áreas vulnerables a esta especie en Kansas, modelamos la distribución potencial de la especie utilizando registros actuales del estado, datos de índices de vegetación tomados remotamente por el Moderate Resolution Imaging Spectrometer (MODIS), y el Genetic Algorithm for Rule-Set Prediction (GARP). Modelos construidos usando sólo las localidades del noreste de Kansas (el origen de la invasión dentro del estado) predijeron consistentemente a las localidades de prueba en otras partes del estado con omisión despreciable. Un análisis adicional usando registros de todos los condados en los cuales se conoce que la especie está presente mostró una predicción similar. Todos los modelos indicaron condiciones apropiadas para la especie en la mayor parte del este y del centro de Kansas, tanto como en áreas ribereñas e irrigadas en la parte occidental del estado. El enfoque presentado aquí posiblemente será de utilidad en la evaluación del potencial de colonización regional de otras especies invasoras recientemente detectadas antes de que otros estudios puedan ser realizados.

Several techniques now exist for modeling the environmental requirements and geographical distribution of a species based on the environmental characteristics of sites of known occurrence (Graham et al., 2004). Such presence-only occurrence data sets often derive from museum and herbarium collections and represent the only data available for many species. Occurrence records are typically employed in conjunction with digital maps that provide data for long-term climatic patterns, topography, potential vegetation, and other environmental variables that commonly influence the macrodistributions of species (Mackey and Lindenmayer, 2001). Using

a Geographic Information System (GIS), a model of the niche requirements of a species (sensu Hutchinson, 1957) in those ecological-environmental dimensions is created and then projected onto geography, thus approximating the potential geographical distribution of the species in the study region (Anderson et al., 2002a).

Such techniques have been applied to the study of invasive species, which constitute a pressing economic and scientific problem affecting diverse sectors of society (Ricciardi et al., 2000). In the past, invasive-species management was largely reactive in nature (Smallwood and Salmon, 1992; Sakai et al., 2001). However,

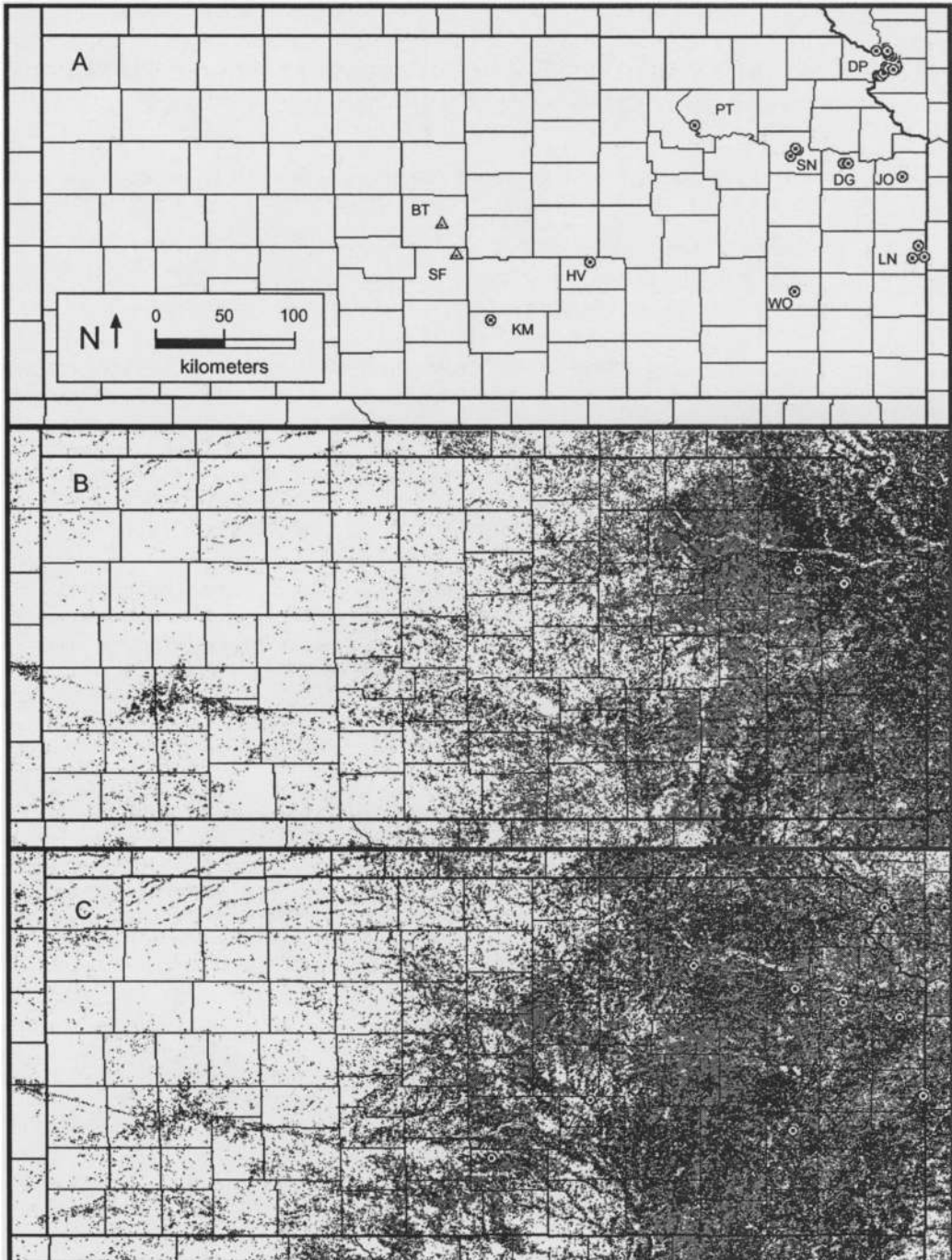


FIG. 1—Known occurrence records and predicted potential distribution of purple loosestrife (*Lythrum salicaria*) along an active invasion front in Kansas. Models were made by using remotely sensed MODIS vegetation-index data and the Genetic Algorithm for Rule-set Prediction (GARP): A) known distributional records (circles) and location of 2 wetlands critical to migratory birds (triangles; Cheyenne Bottoms Wildlife Management Area and Quivira National Wildlife Refuge); B) potential distribution based on the first training data set of locality records from 3 counties in northeastern Kansas (circles in B); C) potential distribution based on records from throughout

building on earlier climatic-matching methods (Johnston, 1924; Chicoine et al., 1985; Panetta and Mitchell, 1991), recent studies have used GIS distributional modeling tools to model the environmental requirements of a species based on occurrence records from its native range; the ecological model is then projected onto other continents to predict which areas are suitable there (Welk et al., 2002; Peterson, 2003; Peterson et al., 2003).

Regrettably, a combination of factors often hinders distributional modeling of newly detected invasive species. First, georeferenced native-range locality data are seldom readily available (Peterson et al., 2003). Second, due to the non-equilibrium circumstances of an active invasion, records from the colonization front alone typically cannot be used along with traditional environmental data (e.g., topographic and climatic variables; Peterson and Cohoon, 1999) to model the regional colonization potential of the species accurately (Welk, 2004). Under such conditions, the environmental characteristics of newly colonized sites often reflect a strongly biased subset of the niche of the species. Furthermore, traditional environmental data (especially climatic variables) typically exhibit strong spatial autocorrelation, and climatic gradients might correspond (spuriously) to the direction of colonization of an invasive, hindering successful modeling of the geographical distribution of the species. Hence, predictions based on these data sources are expected to yield a strong underestimation of the potential distribution of the species (Welk, 2004).

Fortunately, it might be possible to use remotely sensed vegetation-index data in conjunction with locality records from the active colonization front, with the goal of characterizing at least the next steps in the invasion process. Because reflectance parameters present in vegetation-index data sets indicate important phenological patterns, such data hold promise for discriminating between suitable and unsuitable habitat for many species (Egbert et al., 2002).

Satellite data sets have several benefits over traditional climatic coverages, including less time and lower expense of production, finer spatial resolution, and incorporation of up-to-date human modifications of the landscape. Most importantly (and due in large part to these latter 2 factors), they might avoid or at least strongly reduce many of the spatial-autocorrelation pitfalls of traditional climatic coverages and are less likely to exhibit gradients that correspond coincidentally with the direction of colonization of an invasive.

Such predictions would help management personnel assess the extent and location of areas vulnerable to invasion (at the grain of analysis). Because only a few occurrence records are available at the beginning of an invasion by a species, it is unlikely that they represent its full environmental tolerances. However, due to the advantages mentioned above, models based on remotely sensed vegetation-index coverages might be able to characterize important qualities of the sites inhabited to date and yield useful maps of the areas most vulnerable to invasion in the region. The aim is not for management personnel to search exhaustively for established propagules of an invasive species in all map pixels where it is predicted. Rather, assessing the suitability for future spread in a newly invaded region is critical to focusing management efforts and justifying control measures.

*Purple Loosestrife in Kansas*—Purple loosestrife (*Lythrum salicaria*) is an herbaceous, perennial, wetland plant native to Eurasia that has been detrimental as an invasive species in the United States and Canada (Stuckey, 1980; Thompson et al., 1987; Hight et al., 1995; Whitt et al., 1999). The spread of the species through the wetlands of North America began in New England in the early 1800s, and it has now been reported in all 48 of the contiguous United States except Florida (Thompson et al., 1987; Malecki et al., 1993; Blossey et al., 2001). The species is a strong competitor to native species and can dramatically alter wetland habitat (Gaudet and Keddy, 1995;

←

the state (circles in C; one locality record per county). The composite models shown here represent the superimposition of 10 models (gray = 1 to 9 models predicting presence; black = all 10 models). The state appears distorted lengthwise due to projection in geographic coordinates. County abbreviations follow: BT, Barton; DP, Doniphan; DG, Douglas; HV, Harvey; JO, Johnson; KM, Kingman; LN, Linn; PT, Pottawatomie; SN, Shawnee; SF, Stafford; and WO, Woodson.

Twolan-Strutt and Keddy, 1996; but see Morrison, 2002). It often colonizes rapidly via watercourses, especially disturbed areas, such as roadside ditches (Lindgren, 2003). Control efforts have met only limited success, especially because established stands of the species produce large seed banks (Blossey et al., 2001; Denoth and Myers, 2005; Henne et al., 2005).

Purple loosestrife has been recorded recently at several sites in eastern and central Kansas (Fig. 1A). The appearance of this invasive has caused concern for natural systems along rivers, creeks, and other areas with wetlands. Because gradual yet dramatic climatic gradients exist in Kansas (especially a progressive decrease in precipitation from east to west), the suitability of areas in central and western parts of the state for purple loosestrife is not immediately clear (see K uchler, 1974, for a potential vegetation map of Kansas). Special concern exists for 2 of the most critical stopover points for migratory birds in the Great Plains: Cheyenne Bottoms Wildlife Management Area and Quivira National Wildlife Refuge (Zimmerman, 1990; Skagen and Knopf, 1994; W. Scott, pers. comm.), 2 wetland complexes that lie in relatively drier regions to the west of known records of purple loosestrife in the state. A prior study modeling the potential distribution of the species at a continental extent by using traditional climatic data yielded only coarse-grain predictions that were inconclusive for much of Kansas (Welk, 2004). Hence, fine-grain modeling of the potential distribution of the species in the region could indicate which areas are vulnerable to invasion by purple loosestrife and, thus, prove beneficial for developing a management plan.

Our study extends beyond past research linking remotely sensed data to the distribution of purple loosestrife. A previous study used high-resolution photography to map the spatial extent of existing stands of purple loosestrife (Frazier and Moore, 1993). In contrast, we apply GIS-based distributional modeling in a predictive framework. Because the satellite data used here are much coarser than the size of most known patches of the species in Kansas, existing stands of purple loosestrife contribute only insubstantially to the reflectance parameters of a given map pixel, avoiding a potential circularity in the modeling process. Hence, the vegetation-index data are not being used to detect the presence of purple loosestrife and then identify similar areas

in an effort to map current stands of the species. Rather, the present study uses vegetation-index data to characterize the phenological patterns of vegetation patches that are likely vulnerable to invasion by purple loosestrife and then maps such areas. Furthermore, because most wetlands in Kansas are too small to identify using satellite imagery, the current models do not characterize wetland vegetation per se, but rather any larger vegetation assemblage associated with wetland habitats (where purple loosestrife has been found). Hence, purple loosestrife would be expected to invade only wetland habitats within the areas identified as generally suitable by modeling, an issue of scale and grain common in distributional modeling that must be taken into account in interpreting the predictions (Mackey and Lendenmayer, 2001).

**METHODS—GARP and Distributional Modeling**—Recently, several approaches to predictive modeling of the geographical distributions of species have been developed in a Geographic Information System (GIS) environment (Nicholls, 1989; Skov, 2000; Scott et al., 2002). Many modeling techniques require presence-absence data sets, which are only rarely available (see Elith, 2000; Guisan and Zimmerman, 2000). The Genetic Algorithm for Rule-set Prediction (GARP) has shown utility under a wide variety of conditions with presence-only occurrence data, which are more commonly obtainable (e.g., Elith, 2000; Anderson et al., 2003; Peterson and Robbins, 2003; Wiley et al., 2003).

GARP uses environmental data and localities of the occurrence of a species to produce a model of its niche requirements in those ecological-environmental dimensions (Stockwell and Noble, 1992; Stockwell and Peters, 1999). The ecological model is then projected into geographical space to yield a map of the potential distribution of the species in the study region. Due to various possible biotic interactions and contingent historical causes (e.g., failure to disperse), the species might not inhabit all pixels of potential distribution (Anderson et al., 2002a, 2002b). Furthermore, human modifications of the landscape might further alter (usually reduce) the current distribution of a species (Anderson and Mart nez-Meyer, 2004). Rather than a drawback, however, this “overprediction” derives from the niche-based nature of the models and actually allows for synthetic evolutionary and ecological applications, including prediction of the potential distributions of invasive species (Peterson et al., 1999; Anderson et al., 2002a, 2002b; Peterson et al., 2003).

To build the model, GARP evaluates non-random associations between environmental characteristics of localities of known occurrence versus those of the overall study region. It works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection. First, a method is chosen from a set of possibilities (e.g., logistic regression and bioclimatic-envelope rules) and is applied to the data. Then, a rule

is developed and predictive accuracy (*sensu* Stockwell and Peters, 1999) is evaluated via map pixels intrinsically resampled from both the known distribution and from the study region as a whole. The change in predictive accuracy from one iteration to the next is used to evaluate whether a particular rule should be incorporated into the rule-set. As implemented here, the algorithm runs either 1,000 iterations or until the addition of new rules has no appreciable effect on the intrinsic accuracy measure (convergence).

**Data Sources**—Geographic coverages with 250-m  $\times$  250-m pixels comprised the remotely sensed data. We used a time series of 6 Normalized Difference Vegetation Index (NDVI) 16-day composite coverages, drawn from the Moderate Resolution Imaging Spectrometer (MODIS) 250-m Vegetation Index data set (MOD13Q1; <http://edcimswww.cr.usgs.gov/pub/imswelcome/>). They corresponded to composite end-dates of 25 June, 29 September, 16 October, 31 October, and 15 November 2000; and 16 January 2001 (the only dates available at the time). NDVI is calculated as (near-infrared - red)/(near-infrared + red), producing values ranging from -1.0 to 1.0. Because it is generally correlated with photosynthetic activity, NDVI summarizes quantity and vigor of green vegetation (Tucker, 1979; Jackson et al., 1983). High positive NDVI values indicate high densities of green vegetation, while values near or below zero are characteristic of land-cover types that lack green vegetation, such as water, soil, or impervious urban surfaces (e.g., asphalt, concrete). While raw NDVI data sets are most often used for studies focused on vegetation, they can also be used for land-cover mapping, particularly when a time-series of data (several months or more) is available (Loveland et al., 1991; Strahler et al., 1999). In this study, we used raw NDVI values because they are strongly correlated with biomass and leaf area.

Locality data correspond to information provided by the Kansas Department of Agriculture and to specimens housed in the McGregor Herbarium, University of Kansas, Lawrence (Fig. 1A). These records derive from ad hoc surveys and opportunistic collecting. As with most presence-only data sets (especially those associated with museum and herbarium collections), they do not include data regarding the absence of the species from other sites in the region. Nor do we integrate any information regarding abundances (which is generally lacking in such data sources). Localities without GPS readings were georeferenced to the nearest minute using various detailed maps. Thus, the geographical coordinates for many of the localities are accurate only to ca. 1 to 2 km. This lack of precision (less than the resolution of the vegetation-index data) must be taken into account in evaluating the models (see below). Because these current occurrence localities along an active colonization front likely do not represent the full ecological tolerances of the species, the resulting models represent conservative estimates; the true potential distribution for the species might be substantially larger than that predicted here.

**Model Building**—To evaluate the efficacy of modeling based on the current data sets, we conducted a series of validation tests. In distributional modeling, known occurrence localities are commonly divided at random

into 2 groups: a training (calibration) data set used to create the models and a test (validation) data set used to evaluate them (Fielding and Bell, 1997; Peterson et al., 2002). A one-tailed  $\chi^2$  statistic can then be employed to determine whether test localities fall into regions of predicted presence more often than expected by chance, given the proportion of map pixels predicted present by the model. The independent test data set thus provides an extrinsic measure of model significance (better than random prediction). In addition to statistical significance, good models should also show low (or zero) omission of test localities (Peterson et al., 1999; Anderson et al., 2002a, 2003).

Thus, we partitioned the known locality records of purple loosestrife in Kansas by randomly dividing them into training and test data sets. Each time we did this, the training data set approximated the data available early in a regional colonization process (a small number of locality records known from a marginal area of the study region). To do so, we made 5 training data sets by randomly selecting one record from each of 3 northeastern counties with multiple locality records for the species (Doniphan County, 32 records; Douglas County, 2 records; and Shawnee County, 3 records). These counties are in northeastern Kansas, the area of the state first invaded by the species. Each training data set was selected independently; thus, a locality could be chosen in more than one of the 5 training data sets. For each data partition, all other localities in the state (i.e., those not selected for the training data set) formed the respective test data set, which was used only for evaluating the results. Use of only 3 locality records constitutes a minimal training data set for GARP modeling (Stockwell and Peterson, 2002b). However, such a situation would likely be typical during early stages in the invasion by a species. Thus, these analyses provide a difficult, yet realistic, test for GARP in this pilot implementation.

Next, to estimate the potential distribution of the species in Kansas with a larger, less-biased set of localities, we also conducted an analysis based on locality records from all of the 9 counties where purple loosestrife is known in Kansas. For counties with multiple localities, one was randomly selected for inclusion in the training data set (i.e., sampling was stratified by county to minimize bias due to the direction of colonization from the northeast, from which the preponderance of records derive). Hence, the final training set held 9 localities, and all other records were again set aside as a test data set.

Owing to strong stochastic elements characteristic of genetic algorithms, no unique solution is produced by GARP. Hence, production of multiple models is necessary to account for variability among independent runs of model building (Anderson et al., 2002a, 2003; Lim et al., 2002; Wiley et al., 2003). To do so, we made 10 models for each training data set. The 10 models for each training data set were then combined to create a composite model of the potential distribution of the species across the study region. Because all models showed zero intrinsic omission error (omission of training localities), the model-selection procedure outlined in Anderson et al. (2003) was not necessary. The composite model for each training data set likely

TABLE 1—Results of quantitative tests of the predicted potential distribution of purple loosestrife (*Lythrum salicaria*) along an active invasion front in Kansas. Models were made using remotely sensed MODIS vegetation-index data and the Genetic Algorithm for Rule-set Prediction (GARP). Results are presented for each of 5 random data partitions of occurrence localities into training data sets (for making models; one locality from each of 3 northeastern counties) and test data sets (for evaluating models; the remaining 43 records in the state; see Fig. 1). A one-tailed  $\chi^2$  statistic was employed to determine whether test localities fell into regions predicted present in the composite model (pixels predicted by any of the 10 original models made for the respective training data set) more often than expected by chance, given the proportion of map pixels predicted present by the composite model (area). All test localities falling outside the predictions fell within 2 km of pixels of predicted presence and, thus, within the radius of possible error in the coordinates of the locality records.

Data partition	Area	Test localities in/total	$\chi^2$	<i>P</i>	Test localities within 2 km/total
1	0.373	32/43	25.36	<0.001	43/43
2	0.301	26/43	18.86	<0.001	43/43
3	0.355	30/43	22.08	<0.001	43/43
4	0.330	29/43	23.02	<0.001	43/43
5	0.349	30/43	23.08	<0.001	43/43

approximates the potential distribution of the species better than any single model by revealing any consistent signal present in most models, and we use the composite models as our best estimates of the potential distribution of the species in the study region. Pixels predicted by all 10 original models thus reflect a stronger prediction than pixels predicted by only one original model. The analyses yielded a total of 6 composite models, one for each of the 5 training data sets based on localities from northeastern Kansas, and one for the training data set that included one record from each county in the state where the species is known.

Many of the present locality records for purple loosestrife are accurate only to ca. 1 to 2 km (see *Data Sources*). For testing significance in the  $\chi^2$  analyses, we maintained a strict in-out criterion for the test localities (not allowing any radius of error around the locality records). However, in interpreting extrinsic omission rates (omission of test localities), we took into account the accuracy of the locality records by accepting those falling within 2 km of predicted areas.

RESULTS—The composite models made using each of the 5 training data sets of locality records from the 3 northeastern counties indicated areas of potential distribution for purple loosestrife extending throughout much of eastern and central Kansas, as well as in riverine and irrigated areas in the western part of the state (Fig. 1B). These models predicted localities from the corresponding test data sets significantly better than expected at random. Here, 26 to 32 of the 43 test localities fell in pixels predicted by at least one of the 10 original models (26 to 40% strict omission; Table 1; one-tailed  $\chi^2_1 = 18.9$  to 25.4;  $P < 0.001$ ). More importantly, all test localities falling outside areas predicted by the corresponding composite model lay within 2 km of

predicted pixels (and thus within the radius of possible error in the coordinates of the locality records). In fact, for these analyses, all test localities fell either in or within 2 km of pixels predicted by all 10 original models. Thus, the models achieved not only high statistical significance, but also negligible extrinsic omission in areas predicted with high support (Fig. 1B). For each of these data sets, the composite model indicated areas of potential presence for the species at both Cheyenne Bottoms Wildlife Management Area and Quivira National Wildlife Refuge with high support (all 10 original models predicting presence for at least part of each wetland).

The composite model for the training data set based on one locality per county yielded a similar prediction (Fig. 1C). In that analysis, 34 of the 37 test localities lay in pixels predicted by at least one original model, and all test localities fell within 1 km of pixels predicted present by all 10 models. The same general geographical patterns were found as in the preliminary models, but with a moderately stronger prediction in central and western Kansas (Fig. 1C). Again, Cheyenne Bottoms Wildlife Management Area and Quivira National Wildlife Refuge were strongly predicted, with map pixels at these sites indicated by all 10 original models.

DISCUSSION—Because of the non-equilibrium nature of the ongoing colonization of Kansas from the northeast by purple loosestrife, these models made by using the localities currently known in the state likely represent a conservative

(best-case) scenario of the potential distribution of the species in the region (at the present grain of analysis). Nevertheless, even with few and biased locality data (Stockwell and Peterson, 2002*a*, 2002*b*), the models predicted even distant test localities with negligible omission error, indicating surprisingly low overfitting of the training data. Combining climatic or topographical data with vegetation-index coverages might be advantageous in some circumstances. However, due to the low omission error found in the present study, the addition of other variables is apparently not necessary in this system to predict other known occurrences adequately at this grain of analysis.

However, most wetlands in Kansas are too small to identify using satellite imagery. Thus, the grain of our analyses was larger than most of the actual wetland patches purple loosestrife is likely to invade. Due to these factors, of the areas identified as generally suitable by the models, we expect purple loosestrife to invade only wetland habitats present within such pixels (Shamsi and Whitehead, 1977; Stuckey, 1980; Morrison, 2002; Lindgren, 2003; see also Mackey and Lendenmayer, 2001; Pearson et al., 2004).

The composite model resulting from the analysis based on one locality per county represents our most realistic estimate of the potential distribution of the species in the region, albeit probably still conservative at this grain of analysis. This model indicated a high risk of invasion by purple loosestrife for large areas of Kansas (Fig. 1C; pixels predicted by all 10 models represented the areas with highest risk). In addition to the success of the quantitative model-validation tests, the widespread prediction in eastern and central Kansas, but narrow prediction in western Kansas (confined largely to watercourses) makes sense and lends intuitive credibility to the utility of this method. Cheyenne Bottoms Wildlife Management Area and Quivira National Wildlife Refuge are predicted to be at high risk for invasion, with map pixels corresponding to those sites predicted by all models. The general similarity to the previous predictions (which were based on only 3 localities from northeastern Kansas) suggested robustness to the small and geographically biased samples of the first analyses. Areas predicted by the model are likely vulnerable to invasion (because the model indicated their suitability for the species), and the high capacity

for dispersal and strong competitive ability of the species increases the chances that it will reach and colonize these areas (Thompson et al., 1987; Gaudet and Keddy, 1995; Twolan-Strutt and Keddy, 1996). However, the best-case nature of these analyses (which were performed using non-equilibrium records) cautions against concluding that areas not predicted by the model are fully safe. Effective 1 January 2003, a quarantine for purple loosestrife was established in the state by the Kansas Department of Agriculture, and both control and monitoring efforts continue (W. Scott, pers. comm.).

This modeling approach (using remotely sensed data and current records from the region being invaded) might be useful in predicting the regional colonization potential of other recently detected nonnative species before definitive studies can be undertaken. Even given the advantages of remotely sensed vegetation-index data, models based on non-equilibrium occurrence records might still be biased by spurious correlations between the few known nonnative records and patterns of reflectance present in the base coverages. Hence, such models are probably not useful in anticipating the entire geographical potential of an invasion. In addition, species are only likely to inhabit appropriate habitats within the map pixels used in the modeling analyses. Nevertheless, our approach shows promise for minimizing the effects of biases associated with the training localities (which are highly problematic with traditional climatic and topographic data sets; e.g., Welk, 2004) and for anticipating the next steps of an invasion. The vegetation-index data used here are strongly related to phenology and were highly successful in predicting the occurrence of purple loosestrife, a wetland plant. Vegetation-index data might not be as useful for modeling other types of species, and future research should explore the generality of these results for other taxa, as well as the relative utility of vegetation-index data sets from various seasons. In addition, other (i.e., non-MODIS) remotely sensed data sets might be useful for similar applications (e.g., Egbert et al., 2002). Similarly, in addition to use with GARP, this approach could be applied with other presence-only distributional-modeling algorithms (see Elith, 2000; Graham et al., 2004; Elith et al., 2006). Finally, examination of the rule-sets produced by GARP (or other niche-based models produced



by other algorithms) could provide information regarding the conditions required by the species but is beyond the scope of the current paper.

We thank W. Scott (Kansas Department of Agriculture) for supplying occurrence records of purple loosestrife; C. Freeman and C. Morse for access to the University of Kansas McGregor Herbarium; E. Martínez-Meyer, K. McNyset, B. Ray, M. Ortega-Huerta, and B. Wardlow for geographical data; R. Scachetti-Pereira for assistance with GARP; and N. Horning, P. Keddy, C. Raxworthy, and manifold anonymous reviewers for comments. E. Gutiérrez assisted with the Spanish resúmen. This study was supported in part by the Environmental Protection Agency (grant to ATP), the National Science Foundation (National Partnership for Advanced Computational Infrastructure, Earth System Science Thrust; to ATP and SLE); a Roosevelt Postdoctoral Research Fellowship (American Museum of Natural History; to RPA); and funds provided by the Office of the Dean of Science and the Office of the Provost, City College of the City University of New York (RPA).

#### LITERATURE CITED

- ANDERSON, R. P., M. GÓMEZ-LAVERDE, AND A. T. PETERSON. 2002a. Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography* 11:131–141.
- ANDERSON, R. P., D. LEW, AND A. T. PETERSON. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162:211–232.
- ANDERSON, R. P., AND E. MARTÍNEZ-MEYER. 2004. Modeling species' geographic distributions for preliminary conservation assessments: an implementation with the spiny pocket mice (*Heteromys*) of Ecuador. *Biological Conservation* 116:167–179.
- ANDERSON, R. P., A. T. PETERSON, AND M. GÓMEZ-LAVERDE. 2002b. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 98:3–16.
- BLOSSEY, B., L. C. SKINNER, AND J. TAYLOR. 2001. Impact and management of purple loosestrife (*Lythrum salicaria*) in North America. *Biodiversity and Conservation* 10:1787–1807.
- CHICOINE, T. K., P. K. FAY, AND G. A. NIELSEN. 1985. Predicting weed migration from soil and climate maps. *Weed Science* 34:57–61.
- DENOTH, M., AND J. H. MYERS. 2005. Variable success of biological control of *Lythrum salicaria* in British Columbia. *Biological Control* 32:269–279.
- EGBERT, S. L., E. MARTÍNEZ-MEYER, M. ORTEGA-HUERTA, AND A. T. PETERSON. 2002. Use of datasets derived from time-series AVHRR imagery as surrogates for land cover maps in predicting species' distributions. *Proceedings, IEEE International Geoscience and Remote Sensing Symposium*, 4:2337–2339. Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey.
- ELITH, J. 2000. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: S. Ferson and M. Burgman, editors. *Quantitative methods for conservation biology*. Springer, New York. Pages 39–58.
- ELITH, J., C. H. GRAHAM, R. P. ANDERSON, M. DUDÍK, S. FERRIER, A. GUI SAN, R. J. HIJMANS, F. HUETTMMANN, J. LEATHWICK, A. LEHMANN, J. LI, L. G. LOHMANN, B. A. LOISELLE, G. MANION, C. MORITZ, M. NAKAMURA, Y. NAKAZAWA, J. M. OVERTON, A. T. PETERSON, S. J. PHILLIPS, K. RICHARDSON, R. SCACHETTI-PEREIRA, R. SCHAPIRE, J. SOBERÓN, S. WILLIAMS, M. S. WISZ, AND N. E. ZIMMERMAN. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151.
- FIELDING, A. H., AND J. F. BELL. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38–49.
- FRAZIER, B. E., AND B. C. MOORE. 1993. Some tests of film types for remote sensing of purple loosestrife, *Lythrum salicaria*, at low densities. *Wetlands* 13:145–152.
- GAUDET, C. L., AND P. A. KEDDY. 1995. Competitive performance and species distribution in shoreline plant communities: a comparative approach. *Ecology* 76:280–291.
- GRAHAM, C. H., S. FERRIER, F. HUETTMAN, C. MORITZ, AND A. T. PETERSON. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution* 19:497–503.
- GUI SAN, A., AND N. E. ZIMMERMANN. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147–186.
- HENNE, D. C., C. J. LINDGREN, T. S. GABOR, H. R. MURKIN, AND R. E. ROUGHLEY. 2005. An integrated management strategy for the control of purple loosestrife *Lythrum salicaria* L. (Lythraceae) in the Netley-Libau Marsh, southern Manitoba. *Biological Control* 32:319–325.
- HIGHT, S. D., B. BLOSSEY, J. LAING, AND R. DECLERCK-FLOATE. 1995. Establishment of insect biological control agents from Europe against *Lythrum salicaria* in North America. *Biological Control* 24:967–977.
- HUTCHINSON, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22:415–427.
- JACKSON, R. D., P. N. SLATER, AND P. J. PINTER, JR. 1983. Discrimination of growth and water stress in wheat by various vegetation indices through clear and

- turbid atmospheres. *Remote Sensing of Environment* 13:187–208.
- JOHNSTON, T. H. 1924. The relation of climate to the spread of prickly pear. *Transactions and Proceedings of the Royal Society of South Australia* 48: 269–296.
- KÜCHLER, A. W. 1974. A new vegetation map of Kansas. *Ecology* 55:586–604.
- LIM, B. K., A. T. PETERSON, AND M. D. ENGSTROM. 2002. Robustness of ecological niche modeling algorithms for mammals in Guyana. *Biodiversity and Conservation* 11:1237–1246.
- LINDGREN, C. J. 2003. A brief history of purple loosestrife, *Lythrum salicaria*, in Manitoba and its status in 2001. *Canadian Field-Naturalist* 117:100–109.
- LOVELAND, T. R., J. W. MERCHANT, D. O. OHLEN, AND J. F. BROWN. 1991. Development of a land-cover characteristics database for the conterminous U.S. *Photogrammetric Engineering and Remote Sensing* 57:1453–1463.
- MACKEY, B. G., AND D. B. LINDENMAYER. 2001. Towards a hierarchical framework for modeling the spatial distribution of animals. *Journal of Biogeography* 28:1147–1166.
- MALECKI, R. A., B. BLOSSEY, S. D. HIGHT, D. SCHROEDER, L. T. KOK, AND J. R. COULSON. 1993. Biological control of purple loosestrife: a case for using insects as control agents, after rigorous screening, and for integrating release strategies with research. *BioScience* 43:680–686.
- MORRISON, J. A. 2002. Wetland vegetation before and after experimental purple loosestrife removal. *Wetlands* 22:159–169.
- NICHOLLS, A. O. 1989. How to make biological surveys go further with generalized linear models. *Biological Conservation* 50:51–75.
- PANETTA, F. D., AND N. D. MITCHELL. 1991. Homocline analysis and the prediction of weediness. *Weed Research* 31:273–284.
- PEARSON, R. G., T. P. DAWSON, AND C. LIU. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* 27:285–298.
- PETERSON, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78:419–433.
- PETERSON, A. T., AND K. P. COHOON. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling* 117: 159–164.
- PETERSON, A. T., M. PAPES, AND D. A. KLUZA. 2003. Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science* 51:863–868.
- PETERSON, A. T., AND C. R. ROBINS. 2003. Using ecological-niche modeling to predict barred owl invasions with implications for spotted owl conservation. *Conservation Biology* 17:1161–1165.
- PETERSON, A. T., J. SOBERÓN, AND V. SANCHEZ-CORDERO. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285:1265–1267.
- PETERSON, A. T., D. R. B. STOCKWELL, AND D. A. KLUZA. 2002. Distributional prediction based on ecological niche modeling of primary occurrence data. In: J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. *Predicting species occurrences: issues of accuracy and scale*. Island Press, Washington, D.C. Pages 617–623.
- RICCIARDI, A., W. M. STEINER, R. N. MACK, AND D. SIMBERLOFF. 2000. Toward a global information system for invasive species. *BioScience* 50:239–244.
- SAKAI, A. K., F. W. ALLENDORF, J. S. HOLT, D. M. LODGE, J. MOLOFSKY, K. A. WITH, S. BAUGHMAN, R. J. CABIN, J. E. COHEN, N. C. ELLSTRAND, D. E. McCAULEY, P. O'NEIL, I. M. PARKER, J. N. THOMPSON, AND S. G. WELLER. 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics* 32:305–332.
- SCOTT, J. M., P. J. HEGLUND, M. L. MORRISON, J. B. HAUFLE, M. G. RAPHAEL, W. A. WALL, AND F. B. SAMSON, editors. 2002. *Predicting species occurrences: issues of accuracy and scale*. Island Press, Washington, D.C.
- SHAMSI, S. R. A., AND F. H. WHITEHEAD. 1977. Comparative eco-physiology of *Epilobium hirsutum* L. and *Lythrum salicaria* L. III. Mineral nutrition. *Journal of Ecology* 65:55–70.
- SKAGEN, S. K., AND F. L. KNOPF. 1994. Migrating shorebirds and habitat dynamics at a prairie wetland complex. *Wilson Bulletin* 106:91–105.
- SKOV, F. 2000. Potential plant distribution mapping based on climatic similarity. *Taxon* 49:503–515.
- SMALLWOOD, K. S., AND T. P. SALMON. 1992. A rating system for potential exotic bird and mammal pests. *Biological Conservation* 62:149–159.
- STOCKWELL, D., AND D. PETERS. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Science* 13:143–158.
- STOCKWELL, D. R. B., AND I. R. NOBLE. 1992. Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematics and Computers in Simulation* 33:385–390.
- STOCKWELL, D. R. B., AND A. T. PETERSON. 2002a. Controlling bias in biodiversity data. In: J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. *Predicting species occurrences: issues of accuracy and scale*. Island Press, Washington, D.C. Pages 537–546.
- STOCKWELL, D. R. B., AND A. T. PETERSON. 2002b. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148:1–13.

- STRAHLER, A., D. MUCHONEY, J. BORAK, M. FRIEDL, S. GOPAL, E. LAMBIN, AND A. MOODY. 1999. MODIS Land Cover Product Algorithm Theoretical Basis Document (ATBD), version 5.0: MODIS Land Cover and Land Cover Change. Boston University, Boston, Massachusetts.
- STUCKEY, R. L. 1980. Distributional history of *Lythrum salicaria* (purple loosestrife) in North America. *Bartonia* 47:3–20.
- THOMPSON, D. Q., R. L. STUCKEY, AND E. B. THOMPSON. 1987. Spread, impact, and control of purple loosestrife (*Lythrum salicaria*) in North American wetlands. *Fish and Wildlife Report* 2:1–55.
- TUCKER, C. J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8:127–150.
- TWOLAN-STRUTT, L., AND P. A. KEDDY. 1996. Above- and belowground competition intensity in two contrasting wetland plant communities. *Ecology* 77:259–270.
- WELK, E. 2004. Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns: purple loosestrife (*Lythrum salicaria*) in North America. *Ecological Modelling* 179:551–567.
- WELK, E., K. SCHUBERT, AND M. H. HOFFMANN. 2002. Present and potential distribution of invasive garlic mustard (*Alliaria petiolata*) in North America. *Diversity and Distributions* 8:219–233.
- WHITT, M. B., H. H. PRINCE, AND R. R. COX, JR. 1999. Avian use of purple loosestrife dominated habitat relative to other vegetation types in a Lake Huron wetland complex. *Wilson Bulletin* 111:105–114.
- WILEY, E. O., K. M. MCNYSET, A. T. PETERSON, C. R. ROBINS, AND A. M. STEWART. 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography* 16:120–127.
- ZIMMERMAN, J. L. 1990. *Cheyenne Bottoms: wetland in jeopardy*. University Press of Kansas, Lawrence.

*Submitted 7 July 2005. Accepted 30 March 2006.*  
*Associate Editor was Chris Lawver.*